DOCUMENT RESUME

ED 160 611 TM 007 322

AUTHOR Cichon, Donald J.; Olson, George E.
TITLE Psychometrics and Observations: Issues in a Dual Approach to the Study of Classroom Learning Environments.
PUB DATE [Mar 78]
NOTE 35p.; Paper presented at the Annual Meeting of the American Educational Research Association (62nd, Toronto, Ontario, Canada, March 27-31, 1978)

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS *Classroom Environment; *Classroom Observation Techniques; *Course Evaluation; Educational Anthropology; Evaluation Criteria; *Evaluation Methods; Observation; *Questionnaires; Rating Scales; Research Methodology; Senior High Schools; *Student Attitudes; Test Interpretation
IDENTIFIERS Goal Free Evaluation

ABSTRACT

Anthropological methods of classroom observation were combined with the results of student responses to three questionnaires in a study of classroom learning environments. The questionnaires were: Learning Environment Inventory, Class Activities Questionnaire, and ALP (Authenticity, Legitimacy, Productivity) Ethos Instrument. Although the classroom observations were of little aid in the interpretation of individual items or scales, it was judged that apparent discrepancies between classroom observations and questionnaire responses lead to a deeper and more accurate interpretation of the classroom environment. For example, although the class in "Urban and Future Studies" was rated high in "diversity" and "cohesiveness" and discussions were observed to be vigorous, it was rated low in "satisfaction." It was judged that class discussions lacked sufficient goal direction to bring closure or create challenge. This conclusion would not have been reached on the basis of one method alone. Thus combining the two methods may be less costly than extensive educational anthropological observation and more productive than questionnaire data alone. (CTM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED160611

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

D. J. Cichon

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

Psychometrics and Observations: Issues in a Dual Approach to the Study of Classroom Learning Environments

by

Donald J. Cichon

and

George E. Olson

Presented at the
1978 Annual Meeting of the
American Educational Research Association

Research and Development Center
College of Education
Roosevelt University
Chicago, Illinois 60605

TM007 322

Abstract

This paper presents the findings of a classroom learning environment assessment conducted through the joint utilization of paper-pencil instruments and naturalistic observation. Three psychosocial instruments were administered to a class of 25 students whose activities were subsequently observed for ten class periods. The joint interpretation of observational and instrument data led to a more complete understanding of the dynamics of classroom interaction and raised more researchable hypotheses for further study than were obtained from instrument data alone. This dual approach was found to constitute a valuable model for researching the dynamic interplay of classroom events and student perceptions.

In the past few years educational research has seen the rise of anthropological and social psychological observation methods being used within the field. Notable examples of studies which have made systematic use of such methods are The Complexities of an Urban Classroom (Geoffrey and Smith 1968), "Teaching and Learning in English Primary Schools," (Berlack, et.al. 1975), and "Effective Teachers of Eskimo and Indian Students," (Kleinfeld 1975). The increasingly widespread adoption of these methods and the seriousness to which the adoption is attended are evidenced by recent theoretical examinations in AERA presentation of the last three years and in Educational Researcher articles (Lutz and Ramsey 1974; Overhold and Stallings 1976). These treat the methods themselves and their comparisons with the traditional experimental and psychometric paradigms.

Parlett and Hamilton (1976) have published a rather thorough critique of the "agricultural-botanical" paradigm which is traditionally applied to educational evaluation research, concluding that the paradigm "is often a cumbersome and inadequate procedure. (It) falls short of its own tacit claims to be controlled, exact, and unambiguous" (p. 143). They propose and describe the methods for an alternative paradigm, which they label illuminative evaluation. The methods they describe, however, are applicable to educational research beyond strictly evaluation purposes. The model takes account of the wider contexts in which educational programs function, including the actions and perceptions of the participants in the program, the changing nature of a program as it operates, the unintended effects which it may produce, its most critical processes. The model's "primary concern is with description and interpretation rather than measurement and prediction" (p. 144). It is aimed more at hypothesis raising in a thoroughly grounded field setting, rather than at hypothesis testing.

One of the primary characteristics of the anthropological paradigm is that it is "not a standard methodological package but a general research strategy. It aims to be both adaptable and eclectic" (Parlett and Hamilton 1976, p. 147). Because it is adaptable and eclectic, and because of its recent more widespread application in education, many issues regarding its uses and limitations are coming to the fore to be explored and clarified. This article describes a study done on one classroom for the purpose of bringing to light some of these issues in a particular substantive area, that of students' perceptions of their classroom learning environments.

The more particular focus on the study, and thus the scope of the methodological issues discussed, is that psychometric measures were made in the classroom in addition to the anthropological methods. (Both are technically within the bounds of the anthropological methodology, since it is "eclectic." But for clarity in the analysis, we refer to the psychometric and the anthropological components as distinguishable.) The interfacing of these two raises a special set of issues which has not yet been thoroughly explored except in a very limited sense when observations are used to "validate" psychometric instruments.

We wish to explore the interfacing of psychometric and anthropological methods in a case study of students' perceptions of the learning environment of one classroom. Without limiting ourselves to the view of observations as validating or not validating the measured variables, we ask, What do observations add to the measured variables? It is generally agreed that observations may add some deeper insight into the reductionism which results from an instrument's score, but the specific nature of these additions has not been clarified. How do these two methods complement, coincide, and contrast with each other in the raw data they produce and in the interpretations and impli-

cations that can be drawn from them? What are the limitations, powers, and appropriate purposes of each?

It was found in our study that little additional insight into the classroom was gained by the joint consideration of observational and psychometric methods when the individual variables from the psychometric measures were viewed independently of each other. But when the psychometric measures were considered as a basis for developing a grounded theory (Glaser and Strauss 1967) inferred from a combination of data from instruments, observational data, and inferences from both, then power was seen to be gained in understanding the classroom. This understanding was "greater than the sum of the parts" of the individual sources of information about the class. These conclusions are illustrated in the article through the explication of some intended and unintended effects of the curriculum.

The parameters of the discussion in the article consist only of the specific psychometric measures and observational techniques employed in this study, which are detailed below. These specific data collection techniques are, however, commonly used in research studies. Thus, while the results of this study and their implications may not be generalizable to all psychometric and anthropological methods, they do have substantially broad application. The study also had import beyond the methodological issues in that the study of classroom learning environments has received great emphasis in the educational research literature of the past decade. Thus, the methodological issues explored herein should contribute to the progress in the learning environment field.

Background: Classroom Learning Environments

The concept of classroom learning environment, though it has been embodied in a number of research studies in the last decade, is often amorphous. While at least four general categories of meaning of the concept appear in the research studies (Cichon 1975a), their common foundation lies in the theoretical formulation that behavior is a function of the interaction between person and environment as most generally described by Lewin's "life space" (Randhawa and Fu 1973). For Lewin "life space" is a dynamic field in which two interdependent vectors, person and environment, interact to affect behavior (Lewin 1936). The immediate application of this for present purposes is that school learning is a function of the interaction between (1) the student's psychological, physical and social characteristics and (2) the environment of the particular classroom in which he sits.

Two other theoretical formulations have sought to further apply the dual dimensionality of person and environment vectors to educational settings. In one, the Getzels-Thelen "psycho-social model" explicated the interrelationships of the nomothetic dimension (institution, role, and expectation) and the ideographic dimension (individual, personality, and need-disposition) of activity in a social system (the school or classroom) (Getzels and Thelen 1960). In another the dual concepts of personal needs and environmental press developed by Murray (1938) were implemented in institutional studies by Pace and Stern (1958) and others (Stern, Stein, and Bloom 1956).

The recent studies of learning environments begin with the well-accepted, though not always understood, notion that teachers and students recognize that classes are different in terms of not only the students' abilities in regard to the subject, but also in terms of such things as personalities, be-

havior patterns, and interests of the students. Each class seems to have a distinctive "spirit" or "character" which affects how a teacher deals with the class and how the students feel about the class, what they expect of it, how well they perform in it. One conception of classroom learning environment has been operationally defined as: the conglomerate of the students' perceptions of "the relationship of the students to the subject studied, to one another, and to the organizational properties of the class" (Randhawa and Fu 1973, p. 316). This conception was followed in the present study.

Classroom learning environments have significance for study in at least two respects. First, an understanding of classroom learning may be greatly aided by consideration of this environment. Walberg (1970), after a review of the known determinants of student scholastic performance, concluded that most of the variance in learning is accounted for by aptitude and "the environment of learning during instruction" (p. 186). In a series of empirical studies, Anderson and Walberg found clear relationships between a number of classroom environment characteristics and student achievement on both the individual and class levels (Walberg and Anderson 1968; Anderson and Walberg 1968; Anderson 1970). Thus, given that the environment is at least a "mediator" in the classroom learning process for students, then a more complete understanding of the environment is of value.

In a second respect, the characteristics potentially identified in the learning environment are seen as learnings in themselves, especially so in the last few years in the discussions of the "hidden" curriculum. There, the quality of the students' experiences in the classroom is of direct concern for educators. This notion is not new. For Dewey (1938) experience and education were inseparably connected as means and ends, which enjoy in-

separability in general. These notions are employed by Joyce and Weil (1972) in their suggestion that students should be exposed to a variety of learning environments created by different teaching styles or "models."

Anthropological Method in the Setting

The study discussed in this article was the second part of a larger investigation of general methodological concerns in the study of classroom learning environments (Cichon 1975b). For this part, one class from the original sample was chosen for intensive observation by an "outsider." The method used was that of non-participant naturalistic observation as described by Geoffrey and Smith (1968), with some guidance from procedures outlined by Howard Becker (1958). Naturalistic observation is the process in which the observer engages in direct observation of a group in its natural setting, recording everything possible about the setting, indicating separately his interpretation of certain events as especially significant and as being an interpretation. Rather than having preconceived specifications for analyzing the data thus gathered, only a general notion of the questions to be dealt with in the study are in the mind of the observer at the outset. As the data accumulate and patterns begin to emerge from them, constructs for viewing the data are clarified, hypotheses are raised and tested in further observations, and problems are reformulated if need be, until a satisfactory model for explaining the data is obtained, a grounded theory (Glaser and Strauss 1967). The construct(s) for viewing the data and for focussing the observations after the theory begins to be clarified come from such sources as a formal theory which applies to a problem noticed in the data, a review of the literature on a topic of interest, or a theory of the researcher's own device.

Our observations took place continuously over a two week period. The only information about the purpose of the study given to the observer at the outset was that we were studying classroom learning environments and that three instruments to measure aspects of the environment were already given to this class. During this phase the observations were completely open-ended. After one week of observations, the observer was provided with copies of the instruments to study and to help focus further observations. At that time preliminary discussions of the field notes began between the observer (Olson) and the principal investigator (Cichon). In this sense an a priori theory, though vague, was interjected to the observations which followed. The observations were supplemented by frequent informal discussions among the teacher, Olson, and Cichon, a structured interview with the teacher concerning some of the measured aspects of the learning environment in his class (after the observation period), and a discussion with the class (again after the observations) on the same matter. The field notes and class discussion transcripts were analyzed by both the observer and principal investigator. Although interviews, discussions, and observation records were all used as data sources, in discussion of general methods for the remainder of this article the term observations will be used to include all three; in the actual presentation of the findings below the three will be distinguished.

The class in which the observations took place was a high school elective course entitled "Urban and Future Studies," and was composed of 26 juniors and seniors. The observations and the administration of the psychometric measures took place in early May of the school year.

The Psychometric Measures

Three instruments which fit the operational definition of classroom learning environments given above were administered to the students in the class. These instruments were: the Learning Environment Inventory (LEI), and Class Activities Questionnaire (CAQ), and the ALP (acronym for Authenticity, Legitimacy and Productivity).

The LEI was developed in conjunction with evaluation studies of Harvard Project Physics, and derives from "group dimensions" identified by Hemphill and Westie (1950) (Anderson 1973). The inventory consists of 105 statements of possible classroom characterizations to which the students respond on a "strongly agree - "strongly disagree" scale of four points. The scoring and interpretation procedure consists of 15 scales, or group dimensions, each comprising seven items from the inventory. Each student's ratings for each item in a scale are simply added together to obtain his score on that scale. Thus the range of possible scores for each scale is 7 - 28, seven representing strong disagreement that each item in the scale is characteristic of the class, 28 representing strong agreement of same. For class scores, the scale scores for each individual are averaged across the entire class to obtain the class average on each group dimension. The group dimensions measured are: Cohesiveness, Diversity, Formality, Speed, Environment, Cliqueness, Satisfaction, Disorganization, Difficulty, Apathy, Democratic and Competitiveness.

The CAQ was developed in conjunction with the Illinois Gifted Program Evaluation Project. Its stated purpose is "to obtain information concerning cognitive, behavioral, and affective activities. . . students perceived actually occurring in the classroom" (Steele 1969, p. 1). The questionnaire is com-

posed of 27 statements describing possible activities or student roles in the classroom, to which the students respond on a four point "strongly agree" - "strongly disagree" scale. The scoring and interpretation results in 18 "factors" (some being composed of two items): Memory, Translation, Interpretation, Application, Analysis, Synthesis, Evaluation (those seven adapted from the Bloom Taxonomy), Discussion Opportunity, Test/Grade Stress, Lecture, Enthusiasm, Independence, Divergence, Humor, Ideas Valued, Ideas Enjoyed, Teacher Talk, and Student Preparation Time.

The ALP was developed at the University of Chicago as a part of recent investigations of Classroom Ethos (Thelen 1974). Classroom Ethos has been operationalized in terms of three dimensions of classroom experiences: their authenticity, their legitimacy, and their productivity (for which ALP is an acronym). It is assumed that every classroom has some elements of all three dimensions which are of interest, not the degree of presence of any dimension as a whole. That is, an experience in the classroom can be authentic in several different ways, for example, in being cognitively stimulating, in relating to one's past experience, in bringing one to an openness toward others' ideas, or in instilling a desire to participate. It is the specific character of the authenticity that is of interest. In a similar manner, legitimacy and productivity subsume several subelements which enable one to assess the nature of those dimensions in a particular classroom or group of classes. The instrument itself consists of 24 statements each placed on individual cards, describing possible characteristics of learning groups, which the students rank order on a continuum from "most descriptive of this class" to "least descriptive." The content of each item is paraphrased in Table II.

THE ISSUES ILLUSTRATED

Observational Validation of Instruments

Two approaches to the analysis were taken. The first attempted to deal with the three instruments and their components (scales, items, or factors) separately. That is, the class mean ratings for the 15 LEI scales were charted, and the observations (including the field notes, and teacher interview and class discussion transcriptions) were studied in search of consistent patterns of events which suggested the confirmation or contradiction of scales' high or low ratings. The same procedure was applied to the ALP items and the CAQ factors. The underlying assumption was that class means of responses to the instruments might not be entirely valid for "objectively" characterizing the class environment, and that an outside observer might add further evidence to validate students' descriptions of their own class. The results indicate a basic fallacy in the assumption by confirming the distinction between the observers' and the participants' perceptions (interpretations) of events. While at times there is apparent congruence between the observer's data and the students' data, at other times there is not, and therefore in the absence of a consistent pattern of congruence we cannot rely on observations to validate these instruments. A few examples will serve to illustrate.[1]

Table I represents selected LEI scales, along with their class mean ratings and standard deviations, for the class studied. Diversity is rated the highest, and other evidence tends to support the fact that it should be rated high (though it is impossible to say "highest"). In the discussion with the class, when asked why Diversity was rated so high, the first response given was, "Of course it's rated high. We come from so many differ-

ent (ethnic) backgrounds." Most members of the class nodded in agreement with this statement. The tone of the student's response seemed to indicate "It's so obvious, I'm surprised you even asked." When asked for other reasons Diversity was rated so high. no other response was given. Thus there appears to be a rather poignant awareness of at least the ethnic diversity among the members of the class, and a congruence between the observer's interpretation and the students' instrument results.

Table I

Selected LEI Scales, Mean Ratings, and Standard Deviations for the "Observed" Class

Scale	Mean	Standard Deviation	Characteristicness
Diversity	21.4	2.6	Strong
Cohesiveness	19.4	2.9	Strong-moderate
Satisfaction	15.8	3.2	Weak
Goal Direction	17.3	5.0	Moderate-weak

Similarly, Cohesiveness was one of the higher rated items and the field notes on the class discussions (almost verbatim) indicate a spontaneity of contributions and an apparently "comfortable" interchange of disagreements of the kind that generally appear in groups that know each other well. Additionally, the interview with the teacher revealed that earlier on in the semester an effort was made to "socialize the group. . . to get us to understand each other." Several class sessions were spent discussing the "relationships of student to student, of student to teacher," and role-playing simulations took place when problems arose to get the students to understand the problems and solutions.

On the other hand, Satisfaction was rated as rather uncharacteristic of

the class, and this was surprising in terms of the observations. For one, the general tone of class discussions was a vigorous one. The observer noted that roughly two-thirds of the class regularly participated in discussions. For another, student committees were involved in the planning and evaluation of class activities. Also, a variety of speakers were brought into the class to discuss what appeared to be topics (police on juvenile problems and rights, civic organizations concerning minorities, "the future of music" as some examples) relevant to the subject, Urban and Future Studies. It would appear that these would be conditions conducive to the satisfaction of the class. But obviously we have not observed satisfaction itself. We have observed events about which we have tried to make inferences concerning students' feelings. Students' expectations and their evaluation of events such as those described would be at least partially involved in the feeling of satisfaction with the class and these do not appear in observations over such a short period. In this sense, then, the observations by an outsider could not have predicted the students' perceptions of this dimension with any assurance of accuracy.

One final example further illustrates the congruence. Since the range of possible ratings for each LEI scale extends from 7 to 28, 17.5 is the exact middle position, indicating neither agreement nor disagreement that this scale characterized the class. The class mean for Goal Direction was 17.3, lying slightly on the "disagreement" side. Yet the class structure included sessions which attempted to clarify the class goals (not a common activity in classes) and occasional class sessions to "evaluate the progress toward our goals." While several explanations for the co-existence of these activities and the class perception are possible, they would have to be tested as hypo-

theses in further observations, and the present observations would not have predicted, and cannot help explain, this result. The same type of results occurred for the ALP and CAQ.

Thus it appears that the non-participant observer could describe events in the classroom, and, to some degree, interpret the meaning which students ascribe to them. The instruments of study here measure the students' interpretation of the events, not the events themselves, nor an observer's interpretation of them. An outside observer may be of aid in providing a more thorough understanding of a classroom's dynamics by pointing out incongruencies such as those described in the Satisfaction and Goal Direction examples above, and he may raise hypotheses and further test them. But such a process casts little additional light on the elements measured by the instruments themselves. And, the value of the observations in "confirming" the Diversity and Cohesiveness examples above is at least dubious.

Interpretation of Instruments' Similarities

For the second approach to the analysis, we sought to understand some dynamics of a classroom's functioning, using one or more learning environment instruments as a starting point. That is, we used the instruments to collect specific perceptions of the class with some additional information as to the organization of the perceptions (relative salience), and from these hypothesized general principles of operation and effects on students. We will suggest that observations in addition to the pencil-paper measures are valuable in this regard.

Since this is an exploration unguided by specific theory about the relationships among learning environment characteristics, the range of discussion is broad and we will arbitrarily select a few clusters of characteristics in-

dicated by the instruments' results for this class. These will serve as examples of the types of interpretation of a class that may take place. We begin by presenting in Table II the juxtaposition of the results of the three instruments for this observed class. As indicated there, the data from the three instruments for this class show the relative importance given (mean score) to each of the items. Items for the three instruments which are similar in importance and have been interpreted (in the present study) to be similar in content reveal those perceptions which are particularly salient among students. Items which appear similar in content, but have been rated quite differently from one instrument to another represent apparent inconsistencies of perception. We will present first some item clusters that appear to define "themes" of similar student perception, and evidence concerning the saliency and content of these perceptions will be sought in the observations.[2]

From Table II it is observed that across the three instruments the aspects of Diversity (LEI), Independence and Divergence (CAQ), and "thinking new thoughts" (ALP) represent complementary student perceptions. By rating these items as highly characteristic, students indicate that they perceive that class members exhibit different interests, behave independently, are not bound to one solution to a problem, and think some new thoughts of their own. The overlying theme in these items is one of tolerance of individual concerns: individual student interests, individual conclusions, and individual new thoughts.

For the most part, the classroom observations tend to support the high rating given to this theme. Of the class periods observed, three were dominated by student-to-student interaction. Because of the personal nature of many student comments in these classes, the observer inferred that the opportunity existed for expressing individual opinions and ideas. Students were fre

Table II
Characteristics of the "Observed" Class by the Three Instruments

Characteristicness	LEI		ALP			CAQ	
	scale	mean	rank	item	mean	factor	presence in class*
high	Diversity	21.4	1.	L22-issues from larger society	7.3		
			2.	L14-useful learning	7.4	Independence	Much
			3.	A24-desire to contribute	8.8	Divergence	
	Environment	20.2	4.	A8-stimulate new thoughts	9.2		
	Cohesiveness	19.4	5.	L12-clear nature and requirements of tasks	9.7		
			6.	L16-problems which occur in other groups	10.5		
			7.	L7-concentration on significant tasks	10.6		
			8.	L6-exemplify good group process	10.9		
	Cliqueness	18.1					
	Formality	17.9	9.	L21-shared purpose guided group	11.7		
			10.	P15-problem-solving	11.7	Analysis	
			11.	P5-knowledge of progress	11.7	Evaluation	
			12.	P18-diversity aided the group	12.3	Discussion Opportunity	
	Democratic	17.3	13.	P20-contributed special skills	12.7		
	Goal Direction	17.3	14.	L4-good group reasons for activities	12.9	Lecture	
	Difficulty	17.1	15.	A13-clarified previous personal experiences	12.9		
	Disorganization	16.7	16.	P1-group decisions and action on them	13.3		
	Friction	16.5					
	Competitiveness	16.2	17.	P19-helped each other	13.5		
			18.	A11-could be person "I wanted to be"	13.8		
	Satisfaction	15.8	19.	P10-sequence, continuity	15.0	Enthusiasm	None
	Speed	15.6	20.	A2-challenged by others	15.5	Humor	
	Apathy	14.7	21.	A17-excited	16.4		
	Favoritism	14.0	22.	A9-felt like rapping afterward	16.5		
low			23.	A23-felt time passed quickly	17.7		
			24.	P3-accomplished a great deal	18.6		

*See Steele (1969) for the scoring and interpretation procedure.

to disagree with other students and the teacher and often did so openly. Rarely did the teacher state that any particular factual material was to be learned by all or that any opinion was necessarily the right one.

The concern for the individual was also found in comments by the teacher in the interview postdating the observations. He stressed the need for "sensitivity to individual students, their insecurities, their authoritarian backgrounds, and their individual capabilities." These expressions resonate with a concern for the individual, and they support the students' contention that individual concerns are dominant characterizations.

Elements of the class perceived to be rather uncharacteristic or absent were rated low by students. From Table II, it appears that students felt little Satisfaction (LEI), were not very Enthusiastic (CAQ), and did not perceive the class as exciting (ALP). Possibly accompanying this was the ALP perception that "time did not pass quickly." The theme emerging from these perceptions is one of low affect among students.[3]

The observations portray student satisfaction as variable, depending upon the nature of the student activity. In about half of the classes, outside speakers lectured on various topics, and Satisfaction and Enthusiasm were inferred in the observations to be markedly lower in these classes. In the interview, the teacher stated that satisfaction would vary considerably among students, depending upon their individual orientation to the class. "Some demand special attention like they get from other teachers. Some cannot accept the idea of future studies." With respect to Speed of the class, similar variations were observed. During discussions, the class moved swiftly, but during lectures from outside speakers, there was considerable evidence of boredom (glazed expressions, little participation, etc.).

Data from observations and instrument data do not coincide to any large degree with respect to the theme of affect. While observational conclusions have assessed student affect as variable (depending upon the activity) the overriding or mean level of student response concerning affect is quite low. Its variability in the observations suggests that the reasons for its low instrument rating may be quite complex and worth further investigation.

Perceptions which were rated as rather secondary or somewhat doubtful (the middle range of characteristicness in Table II) did not form a common theme across all three instruments, but some items of the LEI and ALP instruments were similar in content. The extent to which students perceived that they knew what the class was trying to do (Goal Direction), the extent to which they perceived a sense of shared purpose to be guiding their behavior (L 21), and having good reasons for what was done (L 4) form a theme.

The observations often refer to the absence of discussion or specification of goal direction during the individual classes. Some examples of the way classes were convened show more clearly what actually took place.

Teacher: (Class #1) "Take time now to re-read the assignment I gave you last week in the futures book. When you are finished, raise your hand so I'll know."

Teacher: (Class #2) "We will show the film today, then will discuss it. . . Christine, would you carry the discussion today?"

Teacher: (Class #3) "Well, I think we have reached the day when two people will give their talk, because they both came on the same day finally. Please take over and wax eloquent."

From these statements, we observed that it is made clear to students what they are to do, and that goals for activity may be quite evident. What appears to

be absent is some direction concerning what to discuss, what content is most appropriate, or in more blatent terms -- what is to be learned.

In the interview, the teacher revealed that periodically considerable attention was given to the assessment of class goals and evaluation of the class purposes with students. He felt that a goal for each class need not be stated, it being subsumed under the overall internalized notions of goal direction engendered by these earlier activities. This bit of history could explain the absence of much "direction" in the observations themselves, except with respect to activity, and could account for the rather low placement of Goal Direction as well. When students perceived Goal Direction, they may have perceived it from the aspect of major objectives, which are only secondarily related to day-to-day tasks. And/or they may have rated it low due to the lack of specific reference to content. In this way, students reacted similar to the observer as to the absence of Goal Direction. (This theme will receive more attention below.)

In summary thus far, it can be seen that the observations' chief function lies in providing more depth of understanding to the perceptions as categorized by the instruments. While the instrument data reveal the class mean level of perception in each category, and in that sense present the predominant perception of the students in the class, this type of characterization alone is a reduction of the complexity of the actual situation of the classroom. The observations explicate some of the aspects lost in the reduction process. From the text of the observations, it appears that the aspects of Diversity, Independence, etc., refer primarily to the content of discussions and not to different modes of activities taking place in the class. In relation to the theme of "affect," the observations point out the particular elements of the

class that may be satisfying or not to students.

The observations serve a value in the case of the cluster interpretation where they failed for individual scales' interpretation, for three reasons. One, quite simply, more pieces of data are available to be connected in developing one's grounded theory in the case of the clusters. Second, and perhaps more important, the nature of the questions asked of the data is broadened in the cluster situation and thus the questions are better matched to the observational data and the methods of their interpretation. And third, the observational data raise questions to be answered in conjunction with the psychometric data, whereas the question is really preformulated when looking at only one specific scale or item from the psychometric results.

Apparent Differences: Interpretive Issues

The observational data are seen to make further contributions to the explication of elements "lost" in the reduction by the instruments when they are used to cast light on apparent inconsistencies in the students' perceptions of the classroom events. We shall label the following cases <u>interpretive issues</u>, for, as will be seen, it cannot be posited that the "inconsistencies" are anywhere but in the mind of the interpreter, and even there they are not in the end inconsistencies.

As the first example of an interpretive issue, recall that the class was strongly characterized as Diverse, Divergent, and Independent. This set of conditions would appear, in our judgment, to be ideal for students to challenge one another. Yet the ALP "challenge" item, A 2, is ranked as very uncharacteristic of this class, rank 20. The field notes provide evidence for much disagreement among class members in apparently vigorous discussion. One therefore raises the question, What else is happening in this classroom that

allows Diversity, Divergence, and Indeperdence to operate and yet leave the students unchallenged by each other?

Several hypotheses are suggested. One, students do not listen to each other and share ideas; they merely express differing opinions until they exhaust themselves. Two, students may listen to each other, but they do not perceive disagreements as challenges. They do not react to an opinion different from their own as suggesting that they question or reorganize their own views or ideas. They may possibly disregard a different opinion from their own as "wrong" or nonsensical. Three, the students merely may not have the facility or techniques available to them to deal more constructively with the diversity that exists in the class.

The second hypothesis suggests an issue of <u>meaning</u>, which the observational evidence does not shed any determinable light on. The first hypothesis is based on the <u>intent</u> of students; i.e., Do they try to listen and share? Again we cannot judge this intent. We are led to the third hypothesis, that of facility in this class. We ask, Is there any evidence to indicate that an attempt to deal constructively with disagreement takes place by the teacher or student discussion leaders in this class?

To answer the question we search for instances of summaries of discussions, attempts to pool the ideas expressed to provide closure, making generalizations from what was said, class discussions which illustrate the complexity of the issues discussed, general principles presented at the outset which are to be illustrated in the discussions, or established content on which to compare and contrast students' own ideas. We find no such instances in the field notes. In each class observed, the bell rang to end the class in the middle of a discussion or lecture. In fact, the apparent vigor of the discus-

sions is so pronounced that no time seems available in class for reflection, a condition indicated as necessary for the educativeness of social inquiry (Thelen, 1960). Thus, we conclude that sufficient mechanism is not provided, as evidenced in the field notes, for creating the "challenge" opportunities in this class.

The hypothesis that no special constructive attempt is made in the class structure to bring the extant Diversity, Divergence, and Independence to closure is further evidenced by the ranking of other ALP items. First, item P 18, diversity aided the group, is ranked 12, the middle position. Thus the diversity existing is not seen by the students as being particularly constructive. Further supporting this general notion are the low rankings of a sense of accomplishment (P 3, ranked 24th), and clarification of previous personal experience (A 13, ranked 15th).

The latter is especially striking in that the field notes are replete with students making use of personal experiences in the discussion. For instance, in one discussion the topic centered around "planning for the future." Individual students illustrated points of the discussion with cases of: "eating when my mother serves me; I don't plan it," "I believe in God and am still afraid of death," "I am preparing for college, even though I'm not sure I'm going." In another discussion concerning students' rights, five instances of alleged violation of class members' rights by school authorities were presented. In a question and answer session with a policewoman, numerous personal experiences with police saturated the session. It appears then that there is ample opportunity for clarification of experiences, but that something other than just their "airing" needs to take place. The entire discussion of the factors absent may also account for the low rating given to the Satisfaction

scale on the LEI.

As a second example of an interpretive "issue," we view some learning environment elements pertaining directly to structural aspects of the class. On the one hand, the class perceives as salient that they clearly understood the nature and requirements of tasks (L 12, ranked 5th), that they concentrated on the significant aspects of tasks (L 7, ranked 7th), and there was a tendency to agree that the class was well organized. On the other hand, there is a tendency toward disagreement that Goal Direction characterized the class, despite the previously mentioned report of the class attending to the goals and evaluating their progress toward them. Thus we ask, Why, in the presence of such clearly structured components, is the Goal Direction not perceived as clear?

One possible explanation might be that the distinction between short-term and long-term objectives is exhibiting itself here. That is, the elements concerning tasks and organization could well be so salient when the discussion topic, or outside speaker's subject, or reading assignment is very clear. But this need only be so on a day-to-day, short-term basis. The wide variety of speakers, of discussion topics, and committee tasks, however, could serve to diffuse the focus of the class. The teacher reports that some 31 speakers were brought into the class during the semester. During the two weeks of observation, speakers came in representing the police, the city Urban Affairs Division, a minority group organization, and music of the future. In another class, a film, Cities and Suburbs, was shown and discussed. One class period involved the discussion of a reading assignment on planning for the future. Another centered around a student report on a community civil rights organization visited by two students. Though the topics are centered around two

themes, Urban and Future studies, the chronological sequence of activities indicated a shift back and forth between the two, and no special transition was evidenced from one class to the next. That this "fragmentation" was evident to the students is suggested by the low ranking of the ALP item of sequence/continuity (P 10, ranked 19th). Thus, we suggest that the nature of short-term objectives and tasks are clear to the students, as reflected in the rankings of L 12 and L 7 and the rating of Disorganization. However, the somewhat fragmented sequencing of the activities leads the students to lose sight of or deemphasize the importance of the long-term goals, reflected in the low rating of Goal Direction.

Another, though related, interpretation is as follows. The fact that the overriding goals and purposes were reviewed in this class throughout the semester suggests that students' perceptions should indeed be oriented to the long-range goals. The question remains, Where do they get the goals for the short-range tasks? Students say that tasks are clear, requirements for activity are sought (L 12), and that significant aspects of tasks are concentrated upon (L 7). All these statements refer to the structure of the activity, the process of going through a task. These do not refer necessarily to content. From the interview data, the teacher stresses the importance of exemplifying for his students and training them in how to communicate with each other, how to act in different situations, how to deal with their authoritarian backgrounds. The observations demonstrate a vigilance to the control of student behavior, a sort of "structured informality." Thus, the items rated as characteristic of the class with

regard to structure refer to activities, and Goal Direction refers to content. That is, a content-process distinction is perceived by these students.

As was the case with the process of explicating "similarities" across the three instruments, the exercise with the interpretive issues leads to the conclusion that the observations have a function in providing more depth of understanding to students' perceptions as expressed through the learning environment instruments. More particularly, for the interpretive issues, the observations aided in organizing descriptive themes, and helped raise and test hypotheses about the influences of class and teacher activities on student perceptions, and about the content of the perceptions.

Summary

The question with which we began was: To what extent does intensive observation of classrooms aid in the interpretation of the instruments' characterizations of them? It is concluded that the observations are of little aid on the level of the individual items, scales or factors. On this level, the difference between the interpretations of events by participants and by an observer was indicated as limiting the nature of the relationships that might be drawn between observations and instrument responses.

On another level, however, the observations are of great value when used to supplement the interpretation of general themes suggested by clusters of elements of the instruments. It was found that when these instruments are used to suggest themes which describe the organization of students' perceptions about various aspects of the classroom's functioning and effects, then the observations help overcome some of the reduction that is imposed by the instrument data alone. That is, broader questions can be raised which more appropriately match the observational data and the methods of interpretation

used herein. There are more data available to develop a grounded theory through raising questions about the different perceptions of the students, suggesting hypotheses about these, suggesting organizational themes in the discrete quantitative data, and enabling initial tests of the formulated hypotheses to be made. Overall, the naturalistic observations lead one to a more complete understanding of the dynamics of interaction among people, activities, and between the two, in the classroom than would be attainable simply from the rather static characterizations obtained from the instruments themselves.

LIGHT ON OTHER AREAS

The clarification of some technical methodological concerns as described above serves to raise questions and to suggest recommendations in areas of research, evaluation, and curriculum study. In terms of general research into classroom processes, naturalistic observations in the anthropological paradigm are known to be costly in money paid to observers, the great deal of time spent in observing a system, and in the energy and time spent in analysis of the data. Their benefit can be determined only by the purposes of the research. From this study we see that if one is studying only those aspects of the learning environment which derive from student perceptions of selected elements in the classroom, then the observations are of little benefit in comparison with their cost. If one wishes to differentiate perceptions related to different kinds of content and processes that are not known to occur in a classroom, or if one wishes to examine the influence of specific classroom processes on the perceptions of students, then the naturalistic observational methods are extremely beneficial. Yet it was seen here that a relatively short period of observation, two weeks, coupled with the administration of psychometric measures

yielded great benefits for less cost than is generally associated with naturalistic observations.

For specific classroom learning environment research, questions about the meaning of the results from instruments like those used here must be considered. To take but one example discussed above, does a moderate level of class satisfaction, quantitatively measured, mean that the students are moderately satisfied with almost all the experiences in the class? Or does it mean that they are sometimes highly satisfied, sometimes dissatisfied, and thus responding in some "average" way? One result or the other makes a difference in understanding that class, but the answer is not known in the common uses of these instruments. Several possible meanings were raised for other items in the discussion above. This is an issue of content validity. In a sense, all such instruments which ask respondents for their perceptions, with these perceptions being the outcome of interest (that is, no projective interpretations are to be made), have high content validity, for perceptions are requested and received, assuming only the honesty of the respondent. But the value of these outcomes (the perceptions) must be judged in terms of their connection to other phenomena. These connections are made either empirically, by systematic variation with another measured trait, or theoretically, by accurate predictions of certain perceptions when other characteristics are known. The former type of connection is that most often made in the research literature on classroom learning environments using the instruments of this study, and these kinds of connections do not contribute to the development of theory with their maximum potential unless there is a better understanding of the meaning of the scaled perception measure. The results of the present study point to the need for further investigation of the content of students' perceptions given on learning environment instruments.

With respect to evaluation, criticism of traditional evaluation methods is growing.[5] One alternative advocated is Scriven's Goal-Free Evaluation (Scriven 1973). Whereas traditional evaluation of programs begins with the program's goals and attempts to measure the extent to which these goals are attained, that is, the intended effects of the program, Scriven points out that "One of the evaluator's most useful contributions may be to reconceptualize the effects, rather than to regurgitate the (program developer's) conception of them" (p. 320). In goal-free evaluation, it is the evaluator's task to also look for effects that the developer did not intend, for "educators often do more good in unexpected directions than in intended ones" (p. 323). Likewise, one who is implementing a program may be accomplishing the goals quite well, but may be also creating some undesirable effects which he is not aware of. In our "debriefing" with the teacher of the class we studied, we found both of the foregoing items to be true. Thus, we recommend our method as a way of implementing a goal-free evaluation.

Related to goal-free evaluation are comments on the study of curriculum. In order to illustrate the values identified concerning the observations' functions, some examples of "similar" and apparently "contradictory" themes, interpretive issues, were presented in this article. The fact that these were presented as examples is important. That is, other interpreters might identify other themes and issues and inquire about their operation in the same class. It was the explication of the method of interfacing the pencil-paper instruments' results and the observations that is seen to be of primary value. Using the same methods one could approach a classroom from the point of view of any of several conceptions of the teaching-learning process and seek an understanding of the ways in which goals are met, and identify

some effects (intended or unintended) of particular curriculum processes. The method and these (or related) instruments may provide a way of operationally testing the "nurturant effects" of the models of teaching suggested by Joyce and Weil (1972).

With regard to our "substantive" findings about the curriculum in this one class, a limitation must be indicated, one which raises further researchable questions. The limitation stems from the question: Can the similar themes and interpretive issues described in the study be seen as variables? That is, to what extent do they vary over time in this one class, and to what extent would they appear in other classes? For example, if Diversity, Divergence, and Independence are always associated with lack of challenge and clarification of students' previous personal experiences, then attempts to explain that situation are futile in terms of one teacher's lack of constructiveness in dealing with the Diversity, Divergence, and Independence, as we suggested above was the case. It may be that teachers generally cannot deal with it because of some underlying cultural or psychological phenomena independent of the classroom.

RETURN TO THE LARGER FIELD

The methodological points discussed in this article are in one sense not at all new to the field of educational research. Any respectable formulation of the scientific method assumes the presence of the basic parts of that which the anthropological researcher will operationalize in individual ways: incisive powers of observation, tough-minded logical reflection, sound theory, and creative conceptualization. But beyond these a set of procedures was developed in the agricultural-botanical paradigm applied to education which included such notions as: objectivity being insured by establishing hypotheses before the

data are collected, the statement of hypotheses in correspondence with statistical tests, strict random sampling procedures, analysis of variance, and probability inferences. For any number of reasons these latter notions have become accepted as the essentials of educational research methodology, and have become something of a liturgy for the conduct of research in education. Such methods may delude the seeker of knowledge about social systems, as the critics mentioned earlier in the article argue. By shifting the focus of methodological discussion in educational research from the agricultural-botanical paradigm to the less clearly defined anthropological one, ideas which were not obvious in the past can be brought to light. With such increased illumination, the distinctions between the two paradigms should perhaps become less clear and less important for the study of classroom social systems.

Notes

This study was supported by a grant to the Roosevelt University College of Education from the Spencer Foundation. Editorial and substantive suggestions by Roger Simon on an earlier version of this paper were very helpful.

1. The justification of comparing observational and student perceptual data is based upon the following reasoning. Test-retest reliabilities over a four week period on the LEI are reported by Anderson (1973) and on the CAQ by Steele (1969). No reliability data are available for the ALP. The reliabilities of the LEI and CAQ are quite high and thus give confidence that the students' responses on the instruments are more of feelings resulting from their accumulated class experiences in fairly consistent patterns of events rather than of specific recent events. Insofar as the events are consistent over time, then those observed are comparable to the events in which the instruments reflect students' perceptions. Also, the observer's reactions to the events and his inferences about student reactions are thus comparable to the students' perceptions, following from the "consistency assumption above.

2. To investigate the validity of comparing class mean scores across instruments with different measurement scales, the intercorrelations (Pearson r) of the items, scales or factors used in each cluster were examined. These correlations were computed from the class mean values of the items, scales and factors in 11 classes. The reporting of the complete matrices is not within

the scope of this article (they are available in Cichon, 1975b), but the median value of the correlations within each cluster of the "instruments' similarities" are as follows:

Cluster I: Diversity, Independence, Divergence, A8 (median r = +0.50)
Cluster II: Satisfaction, Enthusiasm, A17, A23 (median r = +0.55)
Cluster III: Goal Direction, L21, L4 (median r = +0.15).

The median correlations for Clusters I and II are quite high based on an N of 11, and indicate that the items relate quite well with each other across classes and are thus comparable. The median correlation of Cluster III is lower than desirable, though it is positive, and indicates some tentativeness in interpreting that cluster.

3. For this cluster, the time of the year may have influenced its character.

4. To investigate the validity of mean comparisons across instruments in the clusters of "interpretive issues," intercorrelations within clusters from the same data base as above were used again, but in a slightly different manner. For these, one cluster is chosen to oppose another cluster in terms of their high or low characteristicness of the class. Therefore we want the items within clusters to have small or even negative correlations with the "opposite" clusters. The clusters and their median correlations are:

Cluster IV: Diversity, Divergence, Independence
vs (median r = -0.13,
A3, A13, P3, P18 9 of the 12 were negative)

Cluster V: Goal Direction
vs (median r = +0.004)
Disorganization, L7, L12

As before, these correlations give confidence in the mean score comparisons made in this section.

5. In addition to Parlett and Hamilton (1976) referred to earlier on this point, also see Ross and Cronbach (1976).

References

Anderson, G.J. "Effects of Classroom Social Climate on Individual Learning." American Education Research Journal 7 (1970): 135-152.

Anderson, G.J. The Assessment of Learning Environments: A Manual for the Learning Environment Inventory and the My Class Inventory. Halifax: Nova Scotia, Atlantic Institute of Education, 1973.

Anderson, G.J. and Walberg, H.J. "Classroom Climate and Group Learning." International Journal of Educational Sciences 2 (1968): 175-180.

Becker, H.S. "Problems of Inference and Proof in Participant Observation." American Sociological Review 23 (1958): 652-660.

Berlack, H., et.al. "Teaching and Learning in English Primary Schools." School Review 82(2) (1975): 215-244.

Cichon, D.J. Classroom Learning Environments and Their Subject Area Context. Mimeo, 1975a. (Research and Development Center, College of Education, Roosevelt University, Chicago, Illinois.)

Cichon, D.J. A Study of Classroom Learning Environments. Technical Report #4. Research and Development Center, College of Education, Roosevelt University, 1975b.

Dewey, J. Experience and Education. New York: Collier, 1938.

Geoffrey, W. and Smith, L.M. Complexities of An Urban Classroom. New York: Holt, Rinehart, and Winston, 1968.

Getzels, J.W. and Thelen, H.A. "The Classroom Group as a Unique Social System." In The Dynamics of Instructional Groups: Socio-psychological Aspects of Teaching and Learning, 59th Yearbook of the National Society for the Study of Education, edited by N.B. Henry. Chicago: University of Chicago Press, 1960.

Glaser, B. and Strauss, A. Discovery of Grounded Theory. Chicago: Aldine, 1967.

Hemphill, J.K. and Westie, C.M. "The Measurement of Group Dimensions." Journal of Psychology 29 (1950): 325-342.

Joyce, B. and Weil, M. Models of Teaching. Englewood Cliffs, N.J.: Prentice-Hall, 1972.

Kleinfeld, J. "Effective Teachers of Eskimo and Indian Students." School Review 83(2) (1975): 301-344.

Lewin, K. Principles of Topological Psychology. New York: McGraw-Hill, 1936.

Lutz, F.W. and Ramsey, M.A. "The Use of Anthropological Field Methods in Education." Educational Researcher 3 (1974): 5-9.

Murray, H.A. Explorations in Personality. New York: Oxford University Press, 1938.

Overholt, G.E. and Stallings, W.M. "Ethnographic and Experimental Hypotheses in Educational Research." Educational Researcher 5 (1975): 12-14.

Pace, C.R. and Stern, G.G. "An Approach to the Measurement of Psychological Characteristics of College Environments." Journal of Educational Psychology 49 (1958): 269-277.

Parlett, M. and Hamilton, D. "Evaluation as Illumination: A New Approach to the Study of Innovatory Programs." Evaluation Studies 1 (1976): 140-157.

Randhawa, B.S. and Fu, L.L.W. "Assessment and Effect of Some Classroom Environment Variables." Review of Educational Research 43 (1973): 303-322.

Ross, L. and Cronbach, L.J. "Handbook of Evaluation Research: Essay Review." Educational Researcher 5(10) (1976): 9-19.

Scriven, M. "Goal-Free Evaluation." In School Evaluation: the Politics and Processes, edited by E.R. House. Berkeley: McCutchan, 1973.

Steele, J.M. Dimensions of the Class Activities Questionnaire. Illinois Gifted Program Evaluation, Center for Instructional Research and Curriculum Evaluation, University of Illinois, Urbana, Illinois, 1969.

Stern, G.G., Stein, M.J. and Bloom, B.S. Methods in Personality Assessment. Glencoe, Illinois: The Free Press, 1956.

Thelen, H.A. Education and the Human Quest. Chicago: University of Chicago Press, 1960.

Thelen, H.A. The Educational Ethos of the Midwestern High School. Monograph Report, University of Chicago, Department of Education, 1974.

Walberg, H.J. "A Model for Research on Instruction." School Review 78 (1970): 185-200.

Walberg, H.J. and Anderson, G.J. "Classroom Climate and Individual Learning." Journal of Educational Psychology 59 (1968): 415-419.